On the Optimistic Capacity of Arbitrary Channels*

Po-Ning Chen
Dept. of Commun. Engineering
National Chiao-Tung University
HsinChu, Taiwan, R.O.C.
poning@cc.nctu.edu.tw

Fady Alajaji
Dept. of Mathematics & Statistics
Queen's University
Kingston, ON K7L 3N6, Canada
fady@polya.mast.queensu.ca

Abstract — **A formula for the optimistic capacity of arbitrary channels is established. It is shown to equal the supremum, over all input processes, of the input-output zero-sup-information rate. A general expression for optimistic ε-capacity is also provided.**

I. Overview

The conventional definition of channel capacity C [1] requires the existence of reliable block codes for *all sufficiently large blocklengths.* Alternatively, if it is required that reliable codes exist for *infinitely many blocklengths*, a new, more *optimistic* definition of capacity is obtained [1]. This concept of *optimistic* capacity (denoted by $\bar{C}$) has recently been investigated by Verdú *et.al* for arbitrary single-user channels [1, 2]. More specifically, they provide an (additional) *operational* significance for the optimistic capacity by demonstrating that for a given channel, the classical statement of the source-channel separation theorem holds for every source if and only if $C = \bar{C}$ [2]. They also conjecture that a simple expression for $\bar{C}$ does not exist.

In this paper, we answer the latter point by demonstrating that $\bar{C}$ does indeed have a general formula. The key to showing this result is the application of the generalized sup-information rate introduced in [3] to the existing proofs by Verdú and Han [1] of the direct and converse parts of the conventional coding theorem. A general expression for the optimistic ε-capacity is also established.

II. ε-Inf/Sup-Information Rates

Consider an input process $\boldsymbol{X} \stackrel{\triangle}{=} \{X^n = (X_1^{(n)}, \ldots, X_n^{(n)})\}_{n=1}^{\infty}$ [1]. Denote by $\boldsymbol{Y} \stackrel{\triangle}{=} \{Y^n = (Y_1^{(n)}, \ldots, Y_n^{(n)})\}_{n=1}^{\infty}$ the corresponding output process induced by $\boldsymbol{X}$ via the channel $\boldsymbol{W} \stackrel{\triangle}{=} \{W^n = P_{Y^n|X^n} : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}$. In [4, 1], Han and Verdú introduce the notions of inf/sup-information/entropy rates and illustrate the key role these measures play in proving general traditional source/channel coding theorems. The *inf-information rate* $\underline{I}(\boldsymbol{X};\boldsymbol{Y})$ (resp. *sup-information rate* $\bar{I}(\boldsymbol{X};\boldsymbol{Y})$) between processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined in [4] as the *liminf in probability* (resp. *limsup in prob.*) of the sequence of normalized information densities $\frac{1}{n}\, i_{X^nY^n}(X^n;Y^n)$.

Definition 1 (ε-inf/sup-information rates [3]) The *ε-inf-information rate* $\underline{I}_\varepsilon(\boldsymbol{X};\boldsymbol{Y})$ and *ε-sup-information rate* $\bar{I}_\varepsilon(\boldsymbol{X};\boldsymbol{Y})$ between $\boldsymbol{X}$ and $\boldsymbol{Y}$ are defined by

$$\underline{I}_\varepsilon(\boldsymbol{X};\boldsymbol{Y}) \stackrel{\triangle}{=} \sup\{\delta : \bar{i}_{\boldsymbol{X}\boldsymbol{Y}}(\delta) \le \varepsilon\},$$

where $\bar{i}_{\boldsymbol{X}\boldsymbol{Y}}(\delta) \stackrel{\triangle}{=} \limsup_{n\to\infty} Pr\{(1/n) i_{X^nY^n}(X^n;Y^n) \le \delta\}$,

and $$\bar{I}_\varepsilon(\boldsymbol{X};\boldsymbol{Y}) \stackrel{\triangle}{=} \sup\{\delta : \underline{i}_{\boldsymbol{X}\boldsymbol{Y}}(\delta) \le \varepsilon\},$$

where $\underline{i}_{\boldsymbol{X},\boldsymbol{Y}}(\delta) \stackrel{\triangle}{=} \liminf_{n\to\infty} Pr\{(1/n) i_{X^n,Y^n}(X^n;Y^n) \le \delta\}$.

Note that Han and Verdú's inf/sup information rates are special cases of the above quantities: $\underline{I}(\boldsymbol{X},\boldsymbol{Y}) = \underline{I}_0(\boldsymbol{X},\boldsymbol{Y})$ and $\bar{I}(\boldsymbol{X},\boldsymbol{Y}) = \bar{I}_{1^-}(\boldsymbol{X},\boldsymbol{Y})$.

III. Main Results

Definition 2 Given $0 < \varepsilon < 1$, an (n, M, ε) code for channel $\boldsymbol{W}$ has blocklength n, M codewords, and average error probability not larger than ε. $R \ge 0$ is an optimistic ε-achievable rate if, for every $\delta > 0$, there exist, for infinitely many n, (n, M, ε) codes with rate $\frac{\log M}{n} > R - \delta$. The supremum of optimistic ε-achievable rates is called the optimistic ε-capacity, $\bar{C}_\varepsilon$. The *optimistic channel capacity* $\bar{C}$ is defined as the supremum of the rates that are optimistic ε-achievable for all $0 < \varepsilon < 1$.

Theorem 1 (Optimistic channel coding theorem)

$$\bar{C} = \sup_{\boldsymbol{X}} \bar{I}_0(\boldsymbol{X};\boldsymbol{Y}).$$

Theorem 2 (Optimistic ε-capacity) For $0 < \varepsilon < 1$, the optimistic ε-capacity $\bar{C}_\varepsilon$ satisfies

$$\sup_{\boldsymbol{X}} \bar{I}_{\varepsilon^-}(\boldsymbol{X};\boldsymbol{Y}) \le \bar{C}_\varepsilon \le \sup_{\boldsymbol{X}} \bar{I}_\varepsilon(\boldsymbol{X};\boldsymbol{Y}).$$

Observations

- Recall that the general formula for the (pessimistic) capacity is $C = \sup_{\boldsymbol{X}} \underline{I}(\boldsymbol{X};\boldsymbol{Y})$ [1]. It is known that for a DMC, $C = \bar{C}$. However, in general, $\bar{C} \ge C$ since $\bar{I}_0(\boldsymbol{X};\boldsymbol{Y}) \ge \underline{I}(\boldsymbol{X};\boldsymbol{Y})$ [3].
- A simple example of a channel for which $\bar{C} > C$ is as follows. Consider a nonstationary channel $\boldsymbol{W}$ such that at odd time instances $n = 1, 3, \cdots$, W_n is the transition distribution of a BSC with crossover probability 1/2; and at even time instances $n = 2, 4, 6, \cdots$, W_n is the distribution of a BSC with crossover probability 1/4. Then $C = 0$ and $\bar{C} = 1 - h_b(1/4) > 0$.
- In [5], we further illustrate the application of the generalized information measures of [3] by proving an optimistic general source coding theorem.

References

[1] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, IT-40(4):1147-1157, July 1994.

[2] S. Vembu, S. Verdú and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Trans. Info. Theory*, IT-41(1):44-54, January 1995.

[3] P.-N. Chen and F. Alajaji, "Generalization of information measures," *Proc. Int. Symp. Inform. Theory & Applications*, Victoria, Canada, September 1996.

[4] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, IT-39(3):752-772, May 1993.

[5] P.-N. Chen and F. Alajaji, "Optimistic Shannon coding theorems for arbitrary single-user systems," In preparation.

*This work was supported by an ARC grant from Queen's University and by NSERC of Canada.